

## SEQUENCE LISTING

## (1) GENERAL INFORMATION

- (i) APPLICANT: Regents of the University of Minnesota
- (ii) TITLE OF THE INVENTION: ANALYSIS OF ALPHA INTEGRINS  
FOR THE DIAGNOSIS OF DIABETIC NEPHROPATHY
- (iii) NUMBER OF SEQUENCES: 16
- (iv) CORRESPONDENCE ADDRESS:  
(A) ADDRESSEE: Merchant & Gould  
(B) STREET: 3100 Norwest Center  
90 South 7th Street  
(C) CITY: Minneapolis  
(D) STATE: MN  
(E) COUNTRY: US  
(F) ZIP: 55402
- (v) COMPUTER READABLE FORM:  
(A) MEDIUM TYPE: Diskette  
(B) COMPUTER: IBM Compatible  
(C) OPERATING SYSTEM: DOS  
(D) SOFTWARE: FastSEQ Version 1.5
- (vi) CURRENT APPLICATION DATA:  
(A) APPLICATION NUMBER: Unknown  
(B) FILING DATE:  
(C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:  
(A) APPLICATION NUMBER:  
(B) FILING DATE:
- (viii) ATTORNEY/AGENT INFORMATION:  
(A) NAME: Kettelberger, Denise  
(B) REGISTRATION NUMBER: 33,924  
(C) REFERENCE/DOCKET NUMBER: 600.314USWO
- (ix) TELECOMMUNICATION INFORMATION:  
(A) TELEPHONE: 612-332-5300  
(B) TELEFAX: 612-332-9081  
(C) TELEX:

102050-40000050

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 3987 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

```
(ix)  FEATURE:
      (A) NAME/KEY: CDS
      (B) LOCATION: 420...3959
```

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATG GTC CCC AGG CGT CCT GCC AGC CTA GAG GTC ACT GTA GCC TGC ATA 467  
Met Val Pro Arg Arg Pro Ala Ser Leu Glu Val Thr Val Ala Cys Ile  
-28 -25 -20 -15

TGG	CTT	CTC	ACG	GTC	ATC	CTA	GGC	TTC	TGC	GTC	TCC	TTC	AAT	GTT	GAT	515
Trp	Leu	Leu	Thr	Val	Ile	Leu	Gly	Phe	Cys	Val	Ser	Phe	Asn	Val	Asp	
		-10					-5					1				

GTG AAA AAC TCA ATG AGT TTC AGT GGC CCA GTA GAG GAC ATG TTT GGA 563  
Val Lys Asn Ser Met Ser Phe Ser Gly Pro Val Glu Asp Met Phe Gly  
5 10 15 20

TAC	ACT	GTT	CAA	CAA	TAT	GAA	AAC	GAA	GAA	GGC	AAA	TGG	GTT	CTT	ATT	611
Tyr	Thr	Val	Gln	Gln	Tyr	Glu	Asn	Glu	Glu	Gly	Lys	Trp	Val	Leu	Ile	
				25					30					35		

GGT TCT CCT TTA GTT GGC CAA CCC AAA GCA AGA ACT GGA GAT GTC TAT 659  
Gly Ser Pro Leu Val Gly Gln Pro Lys Ala Arg Thr Gly Asp Val Tyr  
40 45 50

AAG TGT CCG GTT GGG AGA GAG AGA GCA ATG CCT TGC GTG AAG TTG GAC 707  
Lys Cys Pro Val Gly Arg Glu Arg Ala Met Pro Cys Val Lys Leu Asp  
55 60 65

[illegible]

GAA	AAA	TTT	GTG	GAG	GAA	ATA	AAA	TCG	ATC	GCA	AGC	GAG	CCC	ACG	GAA	1427
Glu	Lys	Phe	Val	Glu	Glu	Ile	Lys	Ser	Ile	Ala	Ser	Glu	Pro	Thr	Glu	
		295					300					305				
AAG	CAC	TTC	TTC	AAT	GTC	TCG	GAT	GAG	TTG	GCC	CTG	GTC	ACT	ATT	GTT	1475
Lys	His	Phe	Phe	Asn	Val	Ser	Asp	Glu	Leu	Ala	Leu	Val	Thr	Ile	Val	
		310				315					320					
AAA	GCT	CTG	GGA	GAA	AGG	ATA	TTC	GCT	TTG	GAA	GCG	ACA	GCT	GAC	CAG	1523
Lys	Ala	Leu	Gly	Glu	Arg	Ile	Phe	Ala	Leu	Glu	Ala	Thr	Ala	Asp	Gln	
325					330					335					340	
TCA	GCA	GCT	TCA	TTT	GAG	ATG	GAA	ATG	TCT	CAG	ACT	GGC	TTC	AGT	GCT	1571
Ser	Ala	Ala	Ser	Phe	Glu	Met	Glu	Met	Ser	Gln	Thr	Gly	Phe	Ser	Ala	
				345					350					355		
CAC	TAC	TCC	CAG	GAC	TGG	GTC	ATG	CTT	GGA	GCG	GTG	GGA	GCC	TAT	GAC	1619
His	Tyr	Ser	Gln	Asp	Trp	Val	Met	Leu	Gly	Ala	Val	Gly	Ala	Tyr	Asp	
			360					365					370			
TGG	AAC	GGA	ACT	GTG	GTC	ATG	CAG	AAG	GCT	AAC	CAG	ATG	GTC	ATC	CCT	1667
Trp	Asn	Gly	Thr	Val	Val	Met	Gln	Lys	Ala	Asn	Gln	Met	Val	Ile	Pro	
		375					380					385				
CAT	AAC	ACC	ACC	TTT	CAA	ACT	GAG	CCC	GCC	AAG	ATG	AAC	GAG	CCT	CTG	1715
His	Asn	Thr	Thr	Phe	Gln	Thr	Glu	Pro	Ala	Lys	Met	Asn	Glu	Pro	Leu	
		390				395					400					
GCT	TCT	TAT	TTA	GGT	TAC	ACA	GTG	AAC	TCG	GCC	ACC	ATC	CCT	GGA	GAT	1763
Ala	Ser	Tyr	Leu	Gly	Tyr	Thr	Val	Asn	Ser	Ala	Thr	Ile	Pro	Gly	Asp	
405					410					415					420	
GTG	CTC	TAC	ATC	GCT	GGG	CAG	CCT	CGG	TAC	AAT	CAT	ACG	GGC	CAG	GTC	1811
Val	Leu	Tyr	Ile	Ala	Gly	Gln	Pro	Arg	Tyr	Asn	His	Thr	Gly	Gln	Val	
				425					430					435		
GTC	ATC	TAC	AAG	ATG	GAG	GAT	GGG	AAC	ATC	AAC	ATT	CTG	CAG	ACA	CTC	1859
Val	Ile	Tyr	Lys	Met	Glu	Asp	Gly	Asn	Ile	Asn	Ile	Leu	Gln	Thr	Leu	
			440					445					450			
GGC	GGA	GAG	CAG	ATT	GGT	TCC	TAC	TTT	GGT	AGT	GTC	TTA	ACA	ACA	ATT	1907
Gly	Gly	Glu	Gln	Ile	Gly	Ser	Tyr	Phe	Gly	Ser	Val	Leu	Thr	Thr	Ile	
		455					460					465				
GAC	ATC	GAC	AAA	GAT	TCT	TAT	ACT	GAT	CTG	CTT	CTC	GTC	GGG	GCC	CCC	1955
Asp	Ile	Asp	Lys	Asp	Ser	Tyr	Thr	Asp	Leu	Leu	Leu	Val	Gly	Ala	Pro	
		470				475					480					
ATG	TAC	ATG	GGG	ACA	GAG	AAA	GAG	GAA	CAG	GGC	AAG	GTG	TAC	GTG	TAC	2003
Met	Tyr	Met	Gly	Thr	Glu	Lys	Glu	Glu	Gln	Gly	Lys	Val	Tyr	Val	Tyr	
485					490				495						500	
GCT	GTG	AAT														

GGC AGA CCT GCT GCT CAT CCC TGA AGG ATA ATT CAT GCA CGA AAG AAA	2099
Arg Gln Thr Cys Cys Ser Ser Leu Lys Asp Asn Ser Cys Thr Lys Glu	
520 525 530	
AAC AAG AAT GAG CCC TGC GGG GCC CGC TTC GGA ACA GCA ATT GCT GCT	2147
Asn Lys Asn Glu Pro Cys Gly Ala Arg Phe Gly Thr Ala Ile Ala Ala	
535 540 545	
GTA AAA GAC CTC AAC GTG GAT GGA TTT AAT GAC GTC GTG ATT GGA GCT	2195
Val Lys Asp Leu Asn Val Asp Gly Phe Asn Asp Val Val Ile Gly Ala	
550 555 560	
CCG CTG GAA GAT GAC CAC GCA GGA GCT GTG TAC ATT TAT CAT GGC AGT	2243
Pro Leu Glu Asp Asp His Ala Gly Ala Val Tyr Ile Tyr His Gly Ser	
565 570 575 580	
GGC AAG ACC ATA AGG GAG GCG TAT GCA CAA CGC ATT CCA TCA GGT GGG	2291
Gly Lys Thr Ile Arg Glu Ala Tyr Ala Gln Arg Ile Pro Ser Gly Gly	
585 590 595	
GAT GGC AAG ACC CTG AAA TTT TTC GGC CAG TCT ATC CAC GGA GAG ATG	2339
Asp Gly Lys Thr Leu Lys Phe Phe Gly Gln Ser Ile His Gly Glu Met	
600 605 610	
GAT TTA AAT GGT GAC GGT CTG ACT GAC GTG ACC ATT GGA GGC CTT GGT	2387
Asp Leu Asn Gly Asp Gly Leu Thr Asp Val Thr Ile Gly Gly Leu Gly	
615 620 625	
GGA GCA GCC CTC TTC TGG GCC AGA GAT GTG GCT GTA GTT AAA GTG ACC	2435
Gly Ala Ala Leu Phe Trp Ala Arg Asp Val Ala Val Val Lys Val Thr	
630 635 640	
ATG AAT TTT GAA CCC AAT AAA GTG AAT ATT CAA AAG AAA AAC TGC CGT	2483
Met Asn Phe Glu Pro Asn Lys Val Asn Ile Gln Lys Lys Asn Cys Arg	
645 650 655 660	
GTG GAG GGC AAA GAA ACA GTG TGC ATA AAT GCT ACA ATG TGT TTT CAT	2531
Val Glu Gly Lys Glu Thr Val Cys Ile Asn Ala Thr Met Cys Phe His	
665 670 675	
GTG AAA TTA AAG TCT AAA GAG GAC TCA ATT TAC GAG GCT GAT CTG CAG	2579
Val Lys Leu Lys Ser Lys Glu Asp Ser Ile Tyr Glu Ala Asp Leu Gln	
680 685 690	
TAC CGT GTC ACC CTT GAT TCA CTG AGG CAG ATA TCA CGG AGC TTT TTT	2627
Tyr Arg Val Thr Leu Asp Ser Leu Arg Gln Ile Ser Arg Ser Phe Phe	
695 700 705	
TCT GGA ACT CAG GAA AGG AAG ATT CAA AGA AAT ATC ACC GTT CGA GAA	2675
Ser Gly Thr Gln Glu Arg Lys Ile Gln Arg Asn Ile Thr Val Arg Glu	
710 715 720	
TCA GAA TGC ATC AGG CAC TCC TTC TAC ATG TTG GAC AAA CAT GAC TTT	2723
Ser Glu Cys Ile Arg His Ser Phe Tyr Met Leu Asp Lys His Asp Phe	
725 730 735 740	

CAG GAC TCT GTG AGA GTG ACT CTG GAT TTT AAT CTC ACT GAT CCA GAA Gln Asp Ser Val Arg Val Thr Leu Asp Phe Asn Leu Thr Asp Pro Glu 745 750 755	2771
AAT GGT CCT GTA CTT GAT GAC GCT CTG CCA AAC TCA GTC CAC GAA CAC Asn Gly Pro Val Leu Asp Asp Ala Leu Pro Asn Ser Val His Glu His 760 765 770	2819
ATT CCC TTT GCC AAA GAC TGT GGA AAC AAG GAA AGA TGC ATT TCA GAC Ile Pro Phe Ala Lys Asp Cys Gly Asn Lys Glu Arg Cys Ile Ser Asp 775 780 785	2867
CTC ACT CTG AAT GTG TCC ACC ACA GAA AAG AGC CTG CTG ATC GTC AAG Leu Thr Leu Asn Val Ser Thr Thr Glu Lys Ser Leu Leu Ile Val Lys 790 795 800	2915
TCC CAG CAT GAC AAG TTC AAC GTT AGC CTC ACC GTC AAA AAC AAA GGA Ser Gln His Asp Lys Phe Asn Val Ser Leu Thr Val Lys Asn Lys Gly 805 810 815 820	2963
GAC AGT GCG TAC AAC ACC AGG ACA GTG GTG CAG CAT TCA CCA AAT CTG Asp Ser Ala Tyr Asn Thr Arg Thr Val Val Gln His Ser Pro Asn Leu 825 830 835	3011
ATT TTT TCG GGA ATT GAG GAG ATC CAA AAA GAT AGC TGT GAA TCT AAT Ile Phe Ser Gly Ile Glu Glu Ile Gln Lys Asp Ser Cys Glu Ser Asn 840 845 850	3059
CAA AAT ATC ACT TGC AGA GTT GGA TAT CCT TTC CTA AGA GCA GGA GAA Gln Asn Ile Thr Cys Arg Val Gly Tyr Pro Phe Leu Arg Ala Gly Glu 855 860 865	3107
ACG GTT ACC TTC AAA ATA ATA TTC CAG TTT AAC ACA TCC CAT CTC TCG Thr Val Thr Phe Lys Ile Ile Phe Gln Phe Asn Thr Ser His Leu Ser 870 875 880	3155
GAA AAT GCA ATC ATT CAC TTA AGT GCA ACA AGT GAC AGT GAG GAG CCC Glu Asn Ala Ile Ile His Leu Ser Ala Thr Ser Asp Ser Glu Glu Pro 885 890 895 900	3203
CTG GAA TCT CTT AAT GAT AAT GAA GTA AAT ATT TCC ATC CCA GTA AAA Leu Glu Ser Leu Asn Asp Asn Glu Val Asn Ile Ser Ile Pro Val Lys 905 910 915	3251
TAT GAA GTT GGA CTG CAG TTT TAC AGT TCT GCG AGT GAA CAT CAC ATT Tyr Glu Val Gly Leu Gln Phe Tyr Ser Ser Ala Ser Glu His His Ile 920 925 930	3299
TCA GTC GCT GCC AAT GAG ACG ATC CCT GAG TTT ATT AAC TCC ACT GAG Ser Val Ala Ala Asn Glu Thr Ile Pro Glu Phe Ile Asn Ser Thr Glu 935 940 945	3347
GAC ATT GGG AAT GAA ATT AAT GTC TTC TAT ACG ATT AGA AAG AGG GGG Asp Ile Gly Asn Glu Ile Asn Val Phe Tyr Thr Ile Arg Lys Arg Gly 950 955 960	3395

CAT TTC CCA ATG CCA GAA CTT CAG CTG TCA ATT TCA TTC CCC AAT TTG 3443  
 His Phe Pro Met Pro Glu Leu Gln Leu Ser Ile Ser Phe Pro Asn Leu  
 965 970 975 980

ACG GCA GAT GGT TAT CCT GTA CTG TAC CCA ATT GGA TGG TCA TCT TCA 3491  
 Thr Ala Asp Gly Tyr Pro Val Leu Tyr Pro Ile Gly Trp Ser Ser Ser  
 985 990 995

GAT AAT GTG AAC TGT AGA CCC CGG AGC CTT GAG GAC CCC TTT GGC ATC 3539  
 Asp Asn Val Asn Cys Arg Pro Arg Ser Leu Glu Asp Pro Phe Gly Ile  
 1000 1005 1010

AAC TCT GGG AAG AAA ATG ACA ATA TCG AAG TCT GAG GTT CTC AAA AGA 3587  
 Asn Ser Gly Lys Lys Met Thr Ile Ser Lys Ser Glu Val Leu Lys Arg  
 1015 1020 1025

GGC ACA ATC CAG GAC TGC AGT AGT ACG TGT GGA GTT GCC ACC ATC ACG 3635  
 Gly Thr Ile Gln Asp Cys Ser Ser Thr Cys Gly Val Ala Thr Ile Thr  
 1030 1035 1040

TGT AGC CTC CTT CCT TCC GAC CTG AGT CAA GTG AAT GTC TCG CTC CTC 3683  
 Cys Ser Leu Leu Pro Ser Asp Leu Ser Gln Val Asn Val Ser Leu Leu  
 1045 1050 1055 1060

CTG TGG AAA CCG ACT TTC ATA AGA GCA CAT TTT TCC AGC TTA AAC CTT 3731  
 Leu Trp Lys Pro Thr Phe Ile Arg Ala His Phe Ser Ser Leu Asn Leu  
 1065 1070 1075

ACT CTA AGA GGA GAA CTT AAG AGT GAA AAT TCA TCG CTG ACT TTA AGT 3779  
 Thr Leu Arg Gly Glu Leu Lys Ser Glu Asn Ser Ser Leu Thr Leu Ser  
 1080 1085 1090

AGC AGC AAC CGG AAG CGA GAG CTG GCT ATT CAG ATA TCC AAA GAC GGG 3827  
 Ser Ser Asn Arg Lys Arg Glu Leu Ala Ile Gln Ile Ser Lys Asp Gly  
 1095 1100 1105

CTC CCA GGC AGA GTG CCG CTG TGG GTT ATC CTC CTG AGC GCC TTC GCG 3875  
 Leu Pro Gly Arg Val Pro Leu Trp Val Ile Leu Leu Ser Ala Phe Ala  
 1110 1115 1120

GGG CTA CTG CTG CTA ATG CTC CTT ATA TTG GCT CTG TGG AAG ATT GGA 3923  
 Gly Leu Leu Leu Leu Met Leu Leu Ile Leu Ala Leu Trp Lys Ile Gly  
 1125 1130 1135 1140

TTC TTC AAA AGG CCA CTG AAG AAG AAA ATG GAG AAA TGAAAGGTTT 3969  
 Phe Phe Lys Arg Pro Leu Lys Lys Lys Met Glu Lys  
 1145 1150

CATAGAAAAA AAAAAAAAAA 3987

102050-40000000

## (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1180 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE: N-terminal

(vi) ORIGINAL SOURCE:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

Met Val Pro Arg Arg Pro Ala Ser Leu Glu Val Thr Val Ala Cys Ile
-28          -25          -20          -15

Trp Leu Leu Thr Val Ile Leu Gly Phe Cys Val Ser Phe Asn Val Asp
    -10          -5          1

Val Lys Asn Ser Met Ser Phe Ser Gly Pro Val Glu Asp Met Phe Gly
  5          10          15          20

Tyr Thr Val Gln Gln Tyr Glu Asn Glu Glu Gly Lys Trp Val Leu Ile
          25          30          35

Gly Ser Pro Leu Val Gly Gln Pro Lys Ala Arg Thr Gly Asp Val Tyr
          40          45          50

Lys Cys Pro Val Gly Arg Glu Arg Ala Met Pro Cys Val Lys Leu Asp
          55          60          65

Glu Pro Val Asn Thr Ser Ile Pro Asn Val Thr Glu Ile Lys Glu Asn
  70          75          80

Met Thr Phe Gly Ser Thr Leu Val Thr Asn Pro Asn Gly Gly Phe Leu
  85          90          95          100

Ala Cys Gly Pro Leu Tyr Ala Tyr Arg Cys Gly His Leu His Tyr Thr
          105          110          115

Thr Gly Ile Cys Ser Asp Val Ser Pro Thr Phe Gln Val Val Asn Ser
          120          125          130

Phe Ala Pro Val Gln Glu Cys Ser Thr Gln Leu Asp Ile Val Ile Val
          135          140          145

Leu Asp Gly Ser Asn Ser Ile Tyr Pro Trp Glu Ser Val Ile Ala Phe
          150          155          160

Leu Asn Asp Leu Leu Lys Arg Met Asp Ile Gly Pro Lys Gln Thr Gln
          165          170          175          180

Val Gly Ile Val Gln Tyr Gly Glu Asn Val Thr His Glu Phe Asn Leu
          185          190          195

```

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000



Asn Lys Tyr Ser Ser Thr Glu Glu Val Leu Val Ala Ala Asn Lys Ile  
 200 205 210  
 Gly Arg Gln Gly Gly Leu Gln Thr Met Thr Ala Leu Gly Ile Asp Thr  
 215 220 225  
 Ala Arg Lys Glu Ala Phe Thr Glu Ala Arg Gly Ala Arg Arg Gly Val  
 230 235 240  
 Lys Lys Val Met Val Ile Val Thr Asp Gly Glu Ser His Asp Asn Tyr  
 245 250 255 260  
 Arg Leu Lys Gln Val Ile Gln Asp Cys Glu Asp Glu Asn Ile Gln Arg  
 265 270 275  
 Phe Ser Ile Ala Ile Leu Gly His Tyr Asn Arg Gly Asn Leu Ser Thr  
 280 285 290  
 Glu Lys Phe Val Glu Glu Ile Lys Ser Ile Ala Ser Glu Pro Thr Glu  
 295 300 305  
 Lys His Phe Phe Asn Val Ser Asp Glu Leu Ala Leu Val Thr Ile Val  
 310 315 320  
 Lys Ala Leu Gly Glu Arg Ile Phe Ala Leu Glu Ala Thr Ala Asp Gln  
 325 330 335 340  
 Ser Ala Ala Ser Phe Glu Met Glu Met Ser Gln Thr Gly Phe Ser Ala  
 345 350 355  
 His Tyr Ser Gln Asp Trp Val Met Leu Gly Ala Val Gly Ala Tyr Asp  
 360 365 370  
 Trp Asn Gly Thr Val Val Met Gln Lys Ala Asn Gln Met Val Ile Pro  
 375 380 385  
 His Asn Thr Thr Phe Gln Thr Glu Pro Ala Lys Met Asn Glu Pro Leu  
 390 395 400  
 Ala Ser Tyr Leu Gly Tyr Thr Val Asn Ser Ala Thr Ile Pro Gly Asp  
 405 410 415 420  
 Val Leu Tyr Ile Ala Gly Gln Pro Arg Tyr Asn His Thr Gly Gln Val  
 425 430 435  
 Val Ile Tyr Lys Met Glu Asp Gly Asn Ile Asn Ile Leu Gln Thr Leu  
 440 445 450  
 Gly Gly Glu Gln Ile Gly Ser Tyr Phe Gly Ser Val Leu Thr Thr Ile  
 455 460 465  
 Asp Ile Asp Lys Asp Ser Tyr Thr Asp Leu Leu Leu Val Gly Ala Pro  
 470 475 480  
 Met Tyr Met Gly Thr Glu Lys Glu Glu Gln Gly Lys Val Tyr Val Tyr  
 485 490 495 500

050000004-0502001  
 102050-40000000

Ala Val Asn Gln Thr Arg Phe Glu Tyr Gln Met Ser Leu Glu Pro Ile  
505 510 515

Arg Gln Thr Cys Cys Ser Ser Leu Lys Asp Asn Ser Cys Thr Lys Glu  
520 525 530

Asn Lys Asn Glu Pro Cys Gly Ala Arg Phe Gly Thr Ala Ile Ala Ala  
535 540 545

Val Lys Asp Leu Asn Val Asp Gly Phe Asn Asp Val Val Ile Gly Ala  
550 555 560

Pro Leu Glu Asp Asp His Ala Gly Ala Val Tyr Ile Tyr His Gly Ser  
565 570 575 580

Gly Lys Thr Ile Arg Glu Ala Tyr Ala Gln Arg Ile Pro Ser Gly Gly  
585 590 595

Asp Gly Lys Thr Leu Lys Phe Phe Gly Gln Ser Ile His Gly Glu Met  
600 605 610

Asp Leu Asn Gly Asp Gly Leu Thr Asp Val Thr Ile Gly Gly Leu Gly  
615 620 625

Gly Ala Ala Leu Phe Trp Ala Arg Asp Val Ala Val Val Lys Val Thr  
630 635 640

Met Asn Phe Glu Pro Asn Lys Val Asn Ile Gln Lys Lys Asn Cys Arg  
645 650 655 660

Val Glu Gly Lys Glu Thr Val Cys Ile Asn Ala Thr Met Cys Phe His  
665 670 675

Val Lys Leu Lys Ser Lys Glu Asp Ser Ile Tyr Glu Ala Asp Leu Gln  
680 685 690

Tyr Arg Val Thr Leu Asp Ser Leu Arg Gln Ile Ser Arg Ser Phe Phe  
695 700 705

Ser Gly Thr Gln Glu Arg Lys Ile Gln Arg Asn Ile Thr Val Arg Glu  
710 715 720

Ser Glu Cys Ile Arg His Ser Phe Tyr Met Leu Asp Lys His Asp Phe  
725 730 735 740

Gln Asp Ser Val Arg Val Thr Leu Asp Phe Asn Leu Thr Asp Pro Glu  
745 750 755

Asn Gly Pro Val Leu Asp Asp Ala Leu Pro Asn Ser Val His Glu His  
760 765 770

Ile Pro Phe Ala Lys Asp Cys Gly Asn Lys Glu Arg Cys Ile Ser Asp  
775 780 785

Leu Thr Leu Asn Val Ser Thr Thr Glu Lys Ser Leu Leu Ile Val Lys  
790 795 800

09000004-050201  
102050-40000000

Ser Gln His Asp Lys Phe Asn Val Ser Leu Thr Val Lys Asn Lys Gly  
 805 810 815 820  
 Asp Ser Ala Tyr Asn Thr Arg Thr Val Val Gln His Ser Pro Asn Leu  
 825 830 835  
 Ile Phe Ser Gly Ile Glu Glu Ile Gln Lys Asp Ser Cys Glu Ser Asn  
 840 845 850  
 Gln Asn Ile Thr Cys Arg Val Gly Tyr Pro Phe Leu Arg Ala Gly Glu  
 855 860 865  
 Thr Val Thr Phe Lys Ile Ile Phe Gln Phe Asn Thr Ser His Leu Ser  
 870 875 880  
 Glu Asn Ala Ile Ile His Leu Ser Ala Thr Ser Asp Ser Glu Glu Pro  
 885 890 895 900  
 Leu Glu Ser Leu Asn Asp Asn Glu Val Asn Ile Ser Ile Pro Val Lys  
 905 910 915  
 Tyr Glu Val Gly Leu Gln Phe Tyr Ser Ser Ala Ser Glu His His Ile  
 920 925 930  
 Ser Val Ala Ala Asn Glu Thr Ile Pro Glu Phe Ile Asn Ser Thr Glu  
 935 940 945  
 Asp Ile Gly Asn Glu Ile Asn Val Phe Tyr Thr Ile Arg Lys Arg Gly  
 950 955 960  
 His Phe Pro Met Pro Glu Leu Gln Leu Ser Ile Ser Phe Pro Asn Leu  
 965 970 975 980  
 Thr Ala Asp Gly Tyr Pro Val Leu Tyr Pro Ile Gly Trp Ser Ser Ser  
 985 990 995  
 Asp Asn Val Asn Cys Arg Pro Arg Ser Leu Glu Asp Pro Phe Gly Ile  
 1000 1005 1010  
 Asn Ser Gly Lys Lys Met Thr Ile Ser Lys Ser Glu Val Leu Lys Arg  
 1015 1020 1025  
 Gly Thr Ile Gln Asp Cys Ser Ser Thr Cys Gly Val Ala Thr Ile Thr  
 1030 1035 1040  
 Cys Ser Leu Leu Pro Ser Asp Leu Ser Gln Val Asn Val Ser Leu Leu  
 1045 1050 1055 1060  
 Leu Trp Lys Pro Thr Phe Ile Arg Ala His Phe Ser Ser Leu Asn Leu  
 1065 1070 1075  
 Thr Leu Arg Gly Glu Leu Lys Ser Glu Asn Ser Ser Leu Thr Leu Ser  
 1080 1085 1090  
 Ser Ser Asn Arg Lys Arg Glu Leu Ala Ile Gln Ile Ser Lys Asp Gly  
 1095 1100 1105

090000004-050201  
 102050-10000000

Leu Pro Gly Arg Val Pro Leu Trp Val Ile Leu Leu Ser Ala Phe Ala  
 1110 1115 1120

Gly Leu Leu Leu Leu Met Leu Leu Ile Leu Ala Leu Trp Lys Ile Gly  
 1125 1130 1135 1140

Phe Phe Lys Arg Pro Leu Lys Lys Lys Met Glu Lys  
 1145 1150

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5373 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence

(B) LOCATION: 49...3591

(D) OTHER INFORMATION:

(A) NAME/KEY: mat\_peptide

(B) LOCATION: 136

(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GAATTCCTGC AAACCCAGCG CAACTACGGT CCCCCGGTCA GACCCAGG ATG GGG CCA	57
Met Gly Pro	
-29	
GAA CGG ACA GGG GCC GCG CCG CTG CCG CTG CTG CTG GTG TTA GCG CTC	105
Glu Arg Thr Gly Ala Ala Pro Leu Pro Leu Leu Leu Val Leu Ala Leu	
-25 -20 -15	
AGT CAA GGC ATT TTA AAT TGT TGT TTG GCC TAC AAT GTT GGT CTC CCA	153
Ser Gln Gly Ile Leu Asn Cys Cys Leu Ala Tyr Asn Val Gly Leu Pro	
-10 -5 1 5	
GAA GCA AAA ATA TTT TCC GGT CCT TCA AGT GAA CAG TTT GGG TAT GCA	201
Glu Ala Lys Ile Phe Ser Gly Pro Ser Ser Glu Gln Phe Gly Tyr Ala	
10 15 20	
GTG CAG CAG TTT ATA AAT CCA AAA GGC AAC TGG TTA CTG GTT GGT TCA	249
Val Gln Gln Phe Ile Asn Pro Lys Gly Asn Trp Leu Leu Val Gly Ser	
25 30 35	

CCC Pro	TGG Trp	AGT Ser	GGC Gly	TTT Phe	CCT Pro	GAG Glu	AAC Asn	CGA Arg	ATG Met	GGA Gly	GAT Asp	GTG Val	TAT Tyr	AAA Lys	TGT Cys	297
40						45				50						
CCT Pro	GTT Val	GAC Asp	CTA Leu	TCC Ser	ACT Thr	GCC Ala	ACA Thr	TGT Cys	GAA Glu	AAA Lys	CTA Leu	AAT Asn	TTG Leu	CAA Gln	ACT Thr	345
55				60						65				70		
TCA Ser	ACA Thr	AGC Ser	ATT Ile	CCA Pro	AAT Asn	GTT Val	ACT Thr	GAG Glu	ATG Met	AAA Lys	ACC Thr	AAC Asn	ATG Met	AGC Ser	CTC Leu	393
				75				80						85		
GGC Gly	TTG Leu	ATC Ile	CTC Leu	ACC Thr	AGG Arg	AAC Asn	ATG Met	GGA Gly	ACT Thr	GGA Gly	GGT Gly	TTT Phe	CTC Leu	ACA Thr	TGT Cys	441
		90						95				100				
GGT Gly	CCT Pro	CTG Leu	TGG Trp	GCA Ala	CAG Gln	CAA Gln	TGT Cys	GGG Gly	AAT Asn	CAG Gln	TAT Tyr	TAC Tyr	ACA Thr	ACG Thr	GGT Gly	489
		105				110						115				
GTG Val	TGT Cys	TCT Ser	GAC Asp	ATC Ile	AGT Ser	CCT Pro	GAT Asp	TTT Phe	CAG Gln	CTC Leu	TCA Ser	GCC Ala	AGC Ser	TTC Phe	TCA Ser	537
120						125				130						
CCT Pro	GCA Ala	ACT Thr	CAG Gln	CCC Pro	TGC Cys	CCT Pro	TCC Ser	CTC Leu	ATA Ile	GAT Asp	GTT Val	GTG Val	GTT Val	GTG Val	TGT Cys	585
135				140						145						150
GAT Asp	GAA Glu	TCA Ser	AAT Asn	AGT Ser	ATT Ile	TAT Tyr	CCT Pro	TGG Trp	GAT Asp	GCA Ala	GTA Val	AAG Lys	AAT Asn	TTT Phe	TTG Leu	633
				155				160						165		
GAA Glu	AAA Lys	TTT Phe	GTA Val	CAA Gln	GGC Gly	CTT Leu	GAT Asp	ATA Ile	GGC Gly	CCC Pro	ACA Thr	AAG Lys	ACA Thr	CAG Gln	GTG Val	681
		170						175				180				
GGG Gly	TTA Leu	ATT Ile	CAG Gln	TAT Tyr	GCC Ala	AAT Asn	AAT Asn	CCA Pro	AGA Arg	GTT Val	GTG Val	TTT Phe	AAC Asn	TTG Leu	AAC Asn	729
		185				190						195				
ACA Thr	TAT Tyr	AAA Lys	ACC Thr	AAA Lys	GAA Glu	GAA Glu	ATG Met	ATT Ile	GTA Val	GCA Ala	ACA Thr	TCC Ser	CAG Gln	ACA Thr	TCC Ser	777
200						205				210						
CAA Gln	TAT Tyr	GGT Gly	GGG Gly	GAC Asp	CTC Leu	ACA Thr	AAC Asn	ACA Thr	TTC Phe	GGA Gly	GCA Ala	ATT Ile	CAA Gln	TAT Tyr	GCA Ala	825
215				220						225				230		
AGA Arg	AAA Lys	TAT Tyr	GCC Ala	TAT Tyr	TCA Ser	GCA Ala	GCT Ala	TCT Ser	GGT Gly	GGG Gly	CGA Arg	CGA Arg	AGT Ser	GCT Ala	ACG Thr	873
				235				240						245		
AAA Lys	GTA Val	ATG Met	GTA Val	GTT Val	GTA Val	ACT Thr	GAC Asp	GGT Gly	GAA Glu	TCA Ser	CAT His	GAT Asp	GGT Gly	TCA Ser	ATG Met	921
		250						255				260				

TTG AAA GCT GTG ATT GAT CAA TGC AAC CAT GAC AAT ATA CTG AGG TTT	969
Leu Lys Ala Val Ile Asp Gln Cys Asn His Asp Asn Ile Leu Arg Phe	
265 270 275	
GGC ATA GCA GTT CTT GGG TAC TTA AAC AGA AAC GCC CTT GAT ACT AAA	1017
Gly Ile Ala Val Leu Gly Tyr Leu Asn Arg Asn Ala Leu Asp Thr Lys	
280 285 290	
AAT TTA ATA AAA GAA ATA AAA GCG ATC GCT AGT ATT CCA ACA GAA AGA	1065
Asn Leu Ile Lys Glu Ile Lys Ala Ile Ala Ser Ile Pro Thr Glu Arg	
295 300 305 310	
TAC TTT TTC AAT GTG TCT GAT GAA GCA GCT CTA CTA GAA AAG GCT GGG	1113
Tyr Phe Phe Asn Val Ser Asp Glu Ala Ala Leu Leu Glu Lys Ala Gly	
315 320 325	
ACA TTA GGA GAA CAA ATT TTC AGC ATT GAA GGT ACT GTT CAA GGA GGA	1161
Thr Leu Gly Glu Gln Ile Phe Ser Ile Glu Gly Thr Val Gln Gly Gly	
330 335 340	
GAC AAC TTT CAG ATG GAA ATG TCA CAA GTG GGA TTC AGT GCA GAT TAC	1209
Asp Asn Phe Gln Met Glu Met Ser Gln Val Gly Phe Ser Ala Asp Tyr	
345 350 355	
TCT TCT CAA AAT GAT ATT CTG ATG CTG GGT GCA GTG GGA GCT TTT GGC	1257
Ser Ser Gln Asn Asp Ile Leu Met Leu Gly Ala Val Gly Ala Phe Gly	
360 365 370	
TGG AGT GGG ACC ATT GTC CAG AAG ACA TCT CAT GGC CAT TTG ATC TTT	1305
Trp Ser Gly Thr Ile Val Gln Lys Thr Ser His Gly His Leu Ile Phe	
375 380 385 390	
CCT AAA CAA GCC TTT GAC CAA ATT CTG CAG GAC AGA AAT CAC AGT TCA	1353
Pro Lys Gln Ala Phe Asp Gln Ile Leu Gln Asp Arg Asn His Ser Ser	
395 400 405	
TAT TTA GGT TAC TCT GTG GCT GCA ATT TCT ACT GGA GAA AGC ACT CAC	1401
Tyr Leu Gly Tyr Ser Val Ala Ala Ile Ser Thr Gly Glu Ser Thr His	
410 415 420	
TTT GTT GCT GGT GCT CCT CGG GCA AAT TAT ACC GGC CAG ATA GTG CTA	1449
Phe Val Ala Gly Ala Pro Arg Ala Asn Tyr Thr Gly Gln Ile Val Leu	
425 430 435	
TAT AGT GTG AAT GAG AAT GGC AAT ATC ACG GTT ATT CAG GCT CAC CGA	1497
Tyr Ser Val Asn Glu Asn Gly Asn Ile Thr Val Ile Gln Ala His Arg	
440 445 450	
GGT GAC CAG ATT GGC TCC TAT TTT GGT AGT GTG CTG TGT TCA GTT GAT	1545
Gly Asp Gln Ile Gly Ser Tyr Phe Gly Ser Val Leu Cys Ser Val Asp	
455 460 465 470	
GTG GAT AAA GAC ACC ATT ACA GAC GTG CTC TTG GTA GGT GCA CCA ATG	1593
Val Asp Lys Asp Thr Ile Thr Asp Val Leu Leu Val Gly Ala Pro Met	
475 480 485	

TAC	ATG	AGT	GAC	CTA	AAG	AAA	GAG	GAA	GGA	AGA	GTC	TAC	CTG	TTT	ACT	1641
Tyr	Met	Ser	Asp	Leu	Lys	Lys	Glu	Glu	Gly	Arg	Val	Tyr	Leu	Phe	Thr	
			490					495					500			
ATC	AAA	AAG	GGC	ATT	TTG	GGT	CAG	CAC	CAA	TTT	CTT	GAA	GGC	CCC	GAG	1689
Ile	Lys	Lys	Gly	Ile	Leu	Gly	Gln	His	Gln	Phe	Leu	Glu	Gly	Pro	Glu	
		505					510					515				
GGC	ATT	GAA	AAC	ACT	CGA	TTT	GGT	TCA	GCA	ATT	GCA	GCT	CTT	TCA	GAC	1737
Gly	Ile	Glu	Asn	Thr	Arg	Phe	Gly	Ser	Ala	Ile	Ala	Ala	Leu	Ser	Asp	
	520					525					530					
ATC	AAC	ATG	GAT	GGC	TTT	AAT	GAT	GTG	ATT	GTT	GGT	TCA	CCA	CTA	GAA	1785
Ile	Asn	Met	Asp	Gly	Phe	Asn	Asp	Val	Ile	Val	Gly	Ser	Pro	Leu	Glu	
535					540				545						550	
AAT	CAG	AAT	TCT	GGA	GCT	GTA	TAC	ATT	TAC	AAT	GGT	CAT	CAG	GGC	ACT	1833
Asn	Gln	Asn	Ser	Gly	Ala	Val	Tyr	Ile	Tyr	Asn	Gly	His	Gln	Gly	Thr	
				555				560						565		
ATC	CGC	ACA	AAG	TAT	TCC	CAG	AAA	ATC	TTG	GGA	TCC	GAT	GGA	GCC	TTT	1881
Ile	Arg	Thr	Lys	Tyr	Ser	Gln	Lys	Ile	Leu	Gly	Ser	Asp	Gly	Ala	Phe	
			570				576						580			
AGG	AGC	CAT	CTC	CAG	TAC	TTT	GGG	AGG	TCC	TTG	GAT	GGC	TAT	GGA	GAT	1929
Arg	Ser	His	Leu	Gln	Tyr	Phe	Gly	Arg	Ser	Leu	Asp	Gly	Tyr	Gly	Asp	
		585					590					595				
TTA	AAT	GGG	GAT	TCC	ATC	ACC	GAT	GTG	TCT	ATT	GGT	GCC	TTT	GGA	CAA	1977
Leu	Asn	Gly	Asp	Ser	Ile	Thr	Asp	Val	Ser	Ile	Gly	Ala	Phe	Gly	Gln	
	600					605					610					
GTG	GTT	CAA	CTC	TGG	TCA	CAA	AGT	ATT	GCT	GAT	GTA	GCT	ATA	GAA	GCT	2025
Val	Val	Gln	Leu	Trp	Ser	Gln	Ser	Ile	Ala	Asp	Val	Ala	Ile	Glu	Ala	
615					620				625					630		
TCA	TTC	ACA	CCA	GAA	AAA	ATC	ACT	TTG	GTC	AAC	AAG	AAT	GCT	CAG	ATA	2073
Ser	Phe	Thr	Pro	Glu	Lys	Ile	Thr	Leu	Val	Asn	Lys	Asn	Ala	Gln	Ile	
				635				640						645		
ATT	CTC	AAA	CTC	TGC	TTC	AGT	GCA	AAG	TTC	AGA	CCT	ACT	AAG	CAA	AAC	2121
Ile	Leu	Lys	Leu	Cys	Phe	Ser	Ala	Lys	Phe	Arg	Pro	Thr	Lys	Gln	Asn	
			650				655						660			
AAT	CAA	GTG	GCC	ATT	GTA	TAT	AAC	ATC	ACA	CTT	GAT	GCA	GAT	GGA	TTT	2169
Asn	Gln	Val	Ala	Ile	Val	Tyr	Asn	Ile	Thr	Leu	Asp	Ala	Asp	Gly	Phe	
		665					670					675				
TCA	TCC	AGA	GTA	ACC	TCC	AGG	GGG	TTA	TTT	AAA	GAA	AAC	AAT	GAA	AGG	2217
Ser	Ser	Arg	Val	Thr	Ser	Arg	Gly	Leu	Phe	Lys	Glu	Asn	Asn	Glu	Arg	
		680				685					690					
TGC	CTG	CAG	AAG	AAT	ATG	GTA	GTA	AAT	CAA	GCA	CAG	AGT	TGC	CCC	GAG	2265

CAC His	ATC Ile	ATT Ile	TAT Tyr	ATA Ile 725	CAG Gln	GAG Glu	CCC Pro	TCT Ser	GAT Asp 730	GTT Val	GTC Val	AAC Asn	TCT Ser	TTG Leu 735	GAT Asp	2313
TTG Leu	CGT Arg	GTG Val	GAC Asp 740	ATC Ile	AGT Ser	CTG Leu	GAA Glu 745	AAC Asn	CCT Pro	GGC Gly	ACT Thr	AGC Ser	CCT Pro 750	GCC Ala	CTT Leu	2361
GAA Glu	GCC Ala	TAT Tyr 755	TCT Ser	GAG Glu	ACT Thr	GCC Ala	AAG Lys 760	GTC Val	TTC Phe	AGT Ser	ATT Ile 765	CCT Pro	TTC Phe	CAC His	AAA Lys	2409
GAC Asp 760	TGT Cys	GGT Gly	GAG Glu	GAT Asp	GGA Gly	CTT Leu 765	TGC Cys	ATT Ile	TCT Ser	GAT Asp 770	CTA Leu	GTC Val	CTA Leu	GAT Asp	GTC Val	2457
CGA Arg 775	CAA Gln	ATA Ile	CCA Pro	GCT Ala 780	GCT Ala	CAA Gln	GAA Glu	CAA Gln	CCC Pro	TTT Phe 785	ATT Ile	GTC Val	AGC Ser	AAC Asn	CAA Gln 790	2505
AAC Asn	AAA Lys	AGG Arg	TTA Leu 795	ACA Thr	TTT Phe	TCA Ser	GTA Val	ACA Thr	CTG Leu 800	AAA Lys	AAT Asn	AAA Lys	AGG Arg	GAA Glu 805	AGT Ser	2553
GCA Ala	TAC Tyr	AAC Asn 810	ACT Thr	GGA Gly	ATT Ile	GTT Val	GTT Val	GAT Asp 815	TTT Phe	TCA Ser	GAA Glu	AAC Asn	TTG Leu 820	TTT Phe	TTT Phe	2601
GCA Ala	TCA Ser 825	TTC Phe	TCC Ser	CTA Leu	CCG Pro	GTT Val	GAT Asp 830	GGG Gly	ACA Thr	GAA Glu	GTA Val	ACA Thr 835	TGC Cys	CAG Gln	GTG Val	2649
GCT Ala 840	GCA Ala	TCT Ser	CAG Gln	AAG Lys	TCT Ser	GTT Val 845	GCC Ala	TGC Cys	GAT Asp	GTA Val	GGC Gly 850	TAC Tyr	CCT Pro	GCT Ala	TTA Leu	2697
AAG Lys 855	AGA Arg	GAA Glu	CAA Gln	CAG Gln 860	GTG Val	ACT Thr	TTT Phe	ACT Thr	ATT Ile 865	AAC Asn	TTT Phe	GAC Asp	TTC Phe	AAT Asn	CTT Leu 870	2745
CAA Gln	AAC Asn	CTT Leu	CAG Gln 875	AAT Asn	CAG Gln	GCG Ala	TCT Ser	CTC Leu	AGT Ser 880	TTC Phe	CAA Gln	GCC Ala	TTA Leu	AGT Ser 885	GAA Glu	2793
AGC Ser	CAA Gln	GAA Glu	GAA Glu 890	AAC Asn	AAG Lys	GCT Ala	GAT Asp 895	AAT Asn	TTG Leu	GTC Val	AAC Asn	CTC Leu	AAA Lys 900	ATT Ile	CCT Pro	2841
CTC Leu	CTG Leu	TAT Tyr 905	GAT Asp	GCT Ala	GAA Glu	ATT Ile 910	CAC His	TTA Leu	ACA Thr	AGA Arg	TCT Ser	ACC Thr 915	AAC Asn	ATA Ile	AAT Asn	2889
TTT Phe 920	TAT Tyr	GAA Glu	ATC Ile	TCT Ser	TCG Ser	GAT Asp 925	GGG Gly	AAT Asn	GTT Val	CCT Pro	TCA Ser 930	ATC Ile	GTG Val	CAC His	AGT Ser	2937



TTT	GAA	GAT	GTT	GGT	CCA	AAA	TTC	ATC	TTC	TCC	CTG	AAG	GTA	ACA	ACA	2985
Phe	Glu	Asp	Val	Gly	Pro	Lys	Phe	Ile	Phe	Ser	Leu	Lys	Val	Thr	Thr	
935					940					945					950	
GGA	AGT	GTT	CCA	GTA	AGC	ATG	GCA	ACT	GTA	ATC	ATC	CAC	ATC	CCT	CAG	3033
Gly	Ser	Val	Pro	Val	Ser	Met	Ala	Thr	Val	Ile	Ile	His	Ile	Pro	Gln	
				955					960					965		
TAT	ACC	AAA	GAA	AAG	AAC	CCA	CTG	ATG	TAC	CTA	ACT	GGG	GTG	CAA	ACA	3081
Tyr	Thr	Lys	Glu	Lys	Asn	Pro	Leu	Met	Tyr	Leu	Thr	Gly	Val	Gln	Thr	
			970					975					980			
GAC	AAG	GCT	GGT	GAC	ATC	AGT	TGT	AAT	GCA	GAT	ATC	AAT	CCA	CTG	AAA	3129
Asp	Lys	Ala	Gly	Asp	Ile	Ser	Cys	Asn	Ala	Asp	Ile	Asn	Pro	Leu	Lys	
		985					990					995				
ATA	GGA	CAA	ACA	TCT	TCT	TCT	GTA	TCT	TTC	AAA	AGT	GAA	AAT	TTC	AGG	3177
Ile	Gly	Gln	Thr	Ser	Ser	Ser	Val	Ser	Phe	Lys	Ser	Glu	Asn	Phe	Arg	
	1000					1005					1010					
CAC	ACC	AAA	GAA	TTG	AAC	TGC	AGA	ACT	GCT	TCC	TGT	AGT	AAT	GTT	ACC	3225
His	Thr	Lys	Glu	Leu	Asn	Cys	Arg	Thr	Ala	Ser	Cys	Ser	Asn	Val	Thr	
1015					1020					1025					1030	
TGC	TGG	TTG	AAA	GAC	GTT	CAC	ATG	AAA	GGA	GAA	TAC	TTT	GTT	AAT	GTG	3273
Cys	Trp	Leu	Lys	Asp	Val	His	Met	Lys	Gly	Glu	Tyr	Phe	Val	Asn	Val	
				1035					1040				1045			
ACT	ACC	AGA	ATT	TGG	AAC	GGG	ACT	TTC	GCA	TCA	TCA	ACG	TTC	CAG	ACA	3321
Thr	Thr	Arg	Ile	Trp	Asn	Gly	Thr	Phe	Ala	Ser	Ser	Thr	Phe	Gln	Thr	
			1050					1055					1060			
GTA	CAG	CTA	ACG	GCA	GCT	GCA	GAA	ATC	AAC	ACC	TAT	AAC	CCT	GAG	ATA	3369
Val	Gln	Leu	Thr	Ala	Ala	Ala	Glu	Ile	Asn	Thr	Tyr	Asn	Pro	Glu	Ile	
		1065					1070					1075				
TAT	GTG	ATT	GAA	GAT	AAC	ACT	GTT	ACG	ATT	CCC	CTG	ATG	ATA	ATG	AAA	3417
Tyr	Val	Ile	Glu	Asp	Asn	Thr	Val	Thr	Ile	Pro	Leu	Met	Ile	Met	Lys	
	1080					1085					1090					
CCT	GAT	GAG	AAA	GCC	GAA	GTA	CCA	ACA	GGA	GTT	ATA	ATA	GGA	AGT	ATA	3465
Pro	Asp	Glu	Lys	Ala	Glu	Val	Pro	Thr	Gly	Val	Ile	Ile	Gly	Ser	Ile	
1095					1100					1105					1110	
ATT	GCT	GGA	ATC	CTT	TTG	CTG	TTA	GCT	CTG	GTT	GCA	ATT	TTA	TGG	AAG	3513
Ile	Ala	Gly	Ile	Leu	Leu	Leu	Leu	Ala	Leu	Val	Ala	Ile	Leu	Trp	Lys	
				1115					1120					1125		
CTC	GGC	TTC	TTC	AAA	AGA	AAA	TAT	GAA	AAG	ATG	ACC	AAA	AAT	CCA	GAT	3561
Leu	Gly	Phe	Phe	Lys	Arg	Lys	Tyr	Glu	Lys	Met	Thr	Lys	Asn	Pro	Asp	
			1130													

GGGAACCGGC AGCATCCCAG CCAGGGTTTG CTGTTTGCCT GCATGGATTT CTTTTTAAAT 3675  
 CCCATATTTT TTTTATCATG TCGTAGGTAA ACTAACCTGG TATTTTAAGA GAAAACTGCA 3735  
 GGTCAAGTTT GATGAAGAAA TTGTGGGGGG TGGGGGAGGT GCGGGGGGCA GGTAGGGAAA 3795  
 TAATAGGGAA AATACCTATT TTATATGATG GGGGAAAAAA AGTAATCTTT AAACCTGGCTG 3855  
 GCCCAGAGTT TACATTCTAA TTTGCATTGT GTCAGAAACA TGAAATGCTT CCAAGCATGA 3915  
 CAACTTTTAA AGAAAAATAT GATACTCTCA GATTTTAAAG GGGAAAACTG TTCTCTTTAA 3975  
 AATATTTGTC TTTAAACAGC AACTACAGAA GTGGAAGTGC TTGATATGTA AGTACTTCCA 4035  
 CTTGTGTATA TTTTAATGAA TATTGATGTT AACAAGAGGG GAAAACAAAA CACAGGTTTT 4095  
 TTCAATTTAT GCTGCTCATC CAAAGTTGCC ACAGATGATA CTTCCAAGTG ATAATTTTAT 4155  
 TTATAAACTA GGTAAAATTT GTTGTGTTGTT CCTTTTATAC CACGGCTGCC CCTTCCACAC 4215  
 CCCATCTTGC TCTAATGATC AAAACATGCT TGAATAACTG AGCTTAGAGT ATACCTCCTA 4275  
 TATGTCCATT TAAGTTAGGA GAGGGGGCGA TATAGAGACT AAGGCACAAA ATTTTGTTTA 4335  
 AAACCTCAGAA TATAACATTT ATGTAAAATC CCATCTGCTA GAAGCCCATC CTGTGCCAGA 4395  
 GGAAGGAAAA GGAGGAAATT TCCTTTCTCT TTTAGGAGGC ACAACAGTTC TCTTCTAGGA 4455  
 TTTGTTTGGC TGACTGGCAG TAACCTAGTG AATTTTGTAA AGATGAGTAA TTTCTTTGGC 4515  
 AACCTTCCTC CTCCCTTACT GAACCACTCT CCCACCTCCT GGTGGTACCA TTATTATAGA 4575  
 AGCCCTCTAC AGCCTGACTT TCTCTCCAGC GGTCCAAAGT TATCCCCTCC TTTACCCCTC 4635  
 ATCCAAAGTT CCCACTCCTT CAGGACAGCT GCTGTGCATT AGATATTAGG GGGGAAAGTC 4695  
 ATCTGTTTAA TTTACACACT TGCATGAATT ACTGTATATA AACTCCTTAA CTTCAGGGAG 4755  
 CTATTTTCAT TTAGTGCTAA ACAAGTAAGA AAAATAAGCT AGAGTGAATT TCTAAATGTT 4815  
 GGAATGTTAT GGGATGTAAA CAATGTAAAG TAAAACACTC TCAGGATTTC ACCAGAAGTT 4875  
 ACAGATGAGG CACTGGAAAC CACCACCAAA TTAGCAGGTG CACCTTCTGT GGCTGTCTTG 4935  
 TTTCTGAAGT ACTTTTTCTT CCACAAGAGT GAATTTGACC TAGGCAAGTT TGTTCAAAAAG 4995  
 GTAGATCCTG AGATGATTTG GTCAGATTGG GATAAGGCC AGCAATCTGC ATTTTAAACAA 5055  
 GCACCCAGT CACTAGGATG CAGATGGACC ACACTTTGAG AAACACCACC CATTTCTACT 5115  
 TTTTGCACCT TATTTTCTCT GTTCCTGAGC CCCCACATTC TCTAGGAGAA ACTTAGATTA 5175  
 AAATTCACAG AACTACATA TCTAAAGCTT TGACAAGTCC TTGACCTCTA TAAACTTCAG 5235  
 AGTCCTCATT ATAAAATGGG AAGACTGAGC TGGAGTTCAG CAGTGATGCT TTTTAGTTTT 5295  
 AAAAGTCTAT GATCTGATCT GGACTTCCTA TAATACAAAT ACACAATCCT CCAAGAATTT 5355  
 GACTTGGAAG AGGAATTC 5373

## (2) INFORMATION FOR SEQ ID NO:4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1181 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (iii) HYPOTHETICAL: NO

## (iv) ANTISENSE: NO

## (v) FRAGMENT TYPE: internal

## (vi) ORIGINAL SOURCE:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Gly Pro Glu Arg Thr Gly Ala Ala Pro Leu Pro Leu Leu Val  
 -29 -25 -20 -15

Leu Ala Leu Ser Gln Gly Ile Leu Asn Cys Cys Leu Ala Tyr Asn Val  
 -10 -5 1

Gly Leu Pro Glu Ala Lys Ile Phe Ser Gly Pro Ser Ser Glu Gln Phe  
 5 10 15

Gly Tyr Ala Val Gln Gln Phe Ile Asn Pro Lys Gly Asn Trp Leu Leu  
 20 25 30 35  
 Val Gly Ser Pro Trp Ser Gly Phe Pro Glu Asn Arg Met Gly Asp Val  
 40 45 50  
 Tyr Lys Cys Pro Val Asp Leu Ser Thr Ala Thr Cys Glu Lys Leu Asn  
 55 60 65  
 Leu Gln Thr Ser Thr Ser Ile Pro Asn Val Thr Glu Met Lys Thr Asn  
 70 75 80  
 Met Ser Leu Gly Leu Ile Leu Thr Arg Asn Met Gly Thr Gly Gly Phe  
 85 90 95  
 Leu Thr Cys Gly Pro Leu Trp Ala Gln Gln Cys Gly Asn Gln Tyr Tyr  
 100 105 110 115  
 Thr Thr Gly Val Cys Ser Asp Ile Ser Pro Asp Phe Gln Leu Ser Ala  
 120 125 130  
 Ser Phe Ser Pro Ala Thr Gln Pro Cys Pro Ser Leu Ile Asp Val Val  
 135 140 145  
 Val Val Cys Asp Glu Ser Asn Ser Ile Tyr Pro Trp Asp Ala Val Lys  
 150 155 160  
 Asn Phe Leu Glu Lys Phe Val Gln Gly Leu Asp Ile Gly Pro Thr Lys  
 165 170 175  
 Thr Gln Val Gly Leu Ile Gln Tyr Ala Asn Asn Pro Arg Val Val Phe  
 180 185 190 195  
 Asn Leu Asn Thr Tyr Lys Thr Lys Glu Glu Met Ile Val Ala Thr Ser  
 200 205 210  
 Gln Thr Ser Gln Tyr Gly Gly Asp Leu Thr Asn Thr Phe Gly Ala Ile  
 215 220 225  
 Gln Tyr Ala Arg Lys Tyr Ala Tyr Ser Ala Ala Ser Gly Gly Arg Arg  
 230 235 240  
 Ser Ala Thr Lys Val Met Val Val Val Thr Asp Gly Glu Ser His Asp  
 245 250 255  
 Gly Ser Met Leu Lys Ala Val Ile Asp Gln Cys Asn His Asp Asn Ile  
 260 265 270 275  
 Leu Arg Phe Gly Ile Ala Val Leu Gly Tyr Leu Asn Arg Asn Ala Leu  
 280 285 290  
 Asp Thr Lys Asn Leu Ile Lys Glu Ile Lys Ala Ile Ala Ser Ile Pro  
 295 300 305  
 Thr Glu Arg Tyr Phe Phe Asn Val Ser Asp Glu Ala Ala Leu Leu Glu  
 310 315 320

06000004-05004  
 102050-4000060

[illegible]



[illegible]

[illegible]

(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) **FRAGMENT TYPE:**

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

20

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) **FRAGMENT TYPE:**

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

19

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) **FRAGMENT TYPE:**

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

19

## (2) INFORMATION FOR SEQ ID NO:8:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CATCCATGTT GATGTCTG

18

## (2) INFORMATION FOR SEQ ID NO:9:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CATGTGATTC ACCGTCAG

18

## (2) INFORMATION FOR SEQ ID NO:10:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GCATATTGAA TTGCTCCGAA TGTG

24



## (2) INFORMATION FOR SEQ ID NO:11:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GGCGTATGCA CAACGCA

17

## (2) INFORMATION FOR SEQ ID NO:12:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GCGACAGCTG ACCAGTCAGC A

21

## (2) INFORMATION FOR SEQ ID NO:13:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CACTCCTCCA CAGCTCCT

18

T02050-40000000

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 15 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- ```
(ii) MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:
```

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

ACATGTACTC ACTGG

15

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- ```
(ii) MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:
```

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CTCACATGTG GTCCTCTG

18

(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 22 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- ```
(ii) MOLECULE TYPE: cDNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:
```

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GTCCTGTTGA CCTATCCACT GC

22